

NASA Technical Memorandum 104156

1N-37
63065
7-32

ON MARKOV PARAMETERS IN SYSTEM IDENTIFICATION

Minh Phan, Jer-Nan Juang, and Richard W. Longman

(NASA-TM-104156) ON MARKOV PARAMETERS IN
SYSTEM IDENTIFICATION (NASA) 32 p CSCL 20K

N92-13459

G3/39 Unclass
0053065

October 1991



National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665

On Markov Parameters in System Identification

Minh Phan,¹ Jer-Nan Juang,² and Richard W. Longman³

*Spacecraft Dynamics Branch, M.S. 297
NASA Langley Research Center, Hampton, VA 23665*

Abstract

This paper presents a detailed discussion of Markov parameters in system identification. Different forms of input-output representation of linear discrete-time systems are reviewed and discussed. Interpretation of sampled response data as Markov parameters is presented. Relations between the state-space model and particular linear difference models via the Markov parameters are formulated. A generalization of Markov parameters to observer and Kalman filter Markov parameters for system representation and identification is explained. These extended Markov parameters play an important role in providing not only a state-space realization, but also an observer/Kalman filter for the system of interest.

1. Introduction

The basic objective of system identification is to develop a mathematical model of a physical system by observing its input-output relationship. A mathematical model of a system is required for many purposes, e.g., system analysis or controller design. Often in practice, one concerns with linear models since many physical systems can be described by linear or approximately linear equations. For a given linear system, there are several models that can be used to describe the input-output characteristics of that system. Depending on intended application, one type of model may be preferred over another. Since different types of models all describe the same system, they must be related in a fundamental way. The relationship between different model structures can be described in terms of the Markov parameters, which can be used to describe the system input-output map in various ways. Despite the mathematical connotation attached to its name, the Markov parameters are simply the pulse response functions of a discrete system to a unit pulse input. This paper is an overview of Markov parameters commonly encountered in system identification theory.

The outline of this paper is as follows. The paper starts with a review of different ways to describe the input-output relationship of a linear system. This is standard material, which can be found in various texts.¹⁻⁴ The relationship between the Markov parameters and these different descriptions is discussed. The observer equation can also be used to describe the system input-output relation. The relationship between the Markov parameters of an observer and the Markov parameters of the system is presented. This connection is particularly important since it serves as

¹ Engineer, Lockheed Engineering & Sciences Co., Hampton, Virginia. Formerly, National Research Council Fellow, Spacecraft Dynamics Branch.

² Principal Scientist, Spacecraft Dynamics Branch.

³ National Research Council Senior Fellow, Spacecraft Dynamics Branch. Also, Professor of Mechanical Engineering, Columbia University, New York, New York.

the basic foundation for several recently developed system identification methods,⁵⁻¹¹ including the Observer/Kalman filter Identification algorithm (OKID). The Markov parameters that are discussed in this paper are the discrete-time Markov parameters, but physical dynamic systems are described by continuous-time differential equations. The basic relationship between discrete and continuous-time models is described. The process of converting the original second-order continuous-time dynamic equation to a first-order continuous-time state space equation, and then to a first-order discrete-time state space equation is reviewed. The reverse process of converting a discrete-time identified model to a continuous-time state space representation is also described. The Toeplitz and Hankel matrices which are the two important matrices of Markov parameters are explained. The problem of realization from the Markov parameters is then described with the Eigensystem Realization Algorithm (ERA).¹²⁻¹⁵ Many realization techniques require knowledge of the Markov parameters as a starting point. Different ways of determining the Markov parameters or parameters that have the same structure as those of Markov parameters are presented.⁵⁻¹⁸ This process requires either a simple interpretation of the sampled response under different special test inputs, or a more substantial computation procedure to obtain the Markov parameters from experimental data. In connection with this development, particular stochastic linear difference equations for identification of state space models are also presented and discussed. Basic relations between various stochastic models and the Markov parameters describing a linear system and the associated observer/Kalman filter are formulated.

2. Input-Output Representation

Dynamic systems are modeled by continuous-time or discrete-time equations. In continuous-time models, the input-output relations are described by differential equations. In discrete-time models, they are described by difference equations. A close approximation of a continuous-time model can be obtained by a discrete one provided that the sampling rate is sufficiently high. A linear discrete system is commonly described by an n -th order difference equation, the weighting sequence, or a discrete state space model. Relations between the Markov parameters of a system and these representations are described.

2.1 Linear Difference Equation

A n -th order difference equation representation of a linear discrete time-invariant system with q outputs and m inputs takes the general form

$$y(k) + A_1 y(k-1) + \dots + A_n y(k-n) = B_0 u(k) + B_1 u(k-1) + \dots + B_n u(k-n)$$

or

$$y(k) + \sum_{i=1}^n A_i y(k-i) = \sum_{i=0}^n B_i u(k-i) \quad (1)$$

where the input variable is denoted by $u(k)$, the output variable by $y(k)$, k is the time index, and A_i , B_i are constant matrix coefficients of dimensions $q \times q$, and $q \times m$, respectively. If a delay operator operating on a variable $z(k)$ is defined as

$$q^{-1} z(k) = z(k-1)$$

then Eq. (1) can be written in the following form

$$\mathbf{A}(q^{-1})y(k) = \mathbf{B}(q^{-1})u(k) \quad (2)$$

where the polynomials of the delay operators are given as

$$\mathbf{A}(q^{-1}) = I + A_1q^{-1} + A_2q^{-2} + \dots + A_nq^{-n}$$

$$\mathbf{B}(q^{-1}) = B_0 + B_1q^{-1} + B_2q^{-2} + \dots + B_nq^{-n}$$

This model of a linear discrete system is sometimes referred to as a Deterministic Auto-Regressive Moving Average model (DARMA).⁴ This form is commonly used in developing recursive system identification techniques. The above description bears close resemblance to the transfer function representation of a linear discrete system. The transfer function $\mathbf{H}(z)$ is defined to be such that $\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{U}(z)$, where $\mathbf{U}(z)$ and $\mathbf{Y}(z)$ denote the Z-transform of the input and output sequences, respectively. From Eq. (1),

$$\mathbf{H}(z) = (I + A_1z^{-1} + \dots + A_nz^{-n})^{-1}(B_0 + B_1z^{-1} + \dots + B_nz^{-n}) \quad (3)$$

In the single-variable case, A_i , B_i are scalar elements, denoted by a_i , b_i , the transfer function $\mathbf{H}(z)$ is simply

$$\mathbf{H}(z) = \frac{\mathbf{Y}(z)}{\mathbf{U}(z)} = \frac{b_0 + b_1z^{-1} + \dots + b_nz^{-n}}{1 + a_1z^{-1} + \dots + a_nz^{-n}}$$

2.2 The Weighting Sequence Description

For continuous-time systems governed by linear differential equations, if the response of the system to a Dirac delta impulse is known, its response to any other general input can be determined by the convolution integral. Similarly, for discrete-time systems, analogous to the Dirac impulse function is the Kronecker unit pulse sequence defined as

$$\delta(k) = \begin{cases} 1 & , k = 0 \\ 0 & , k \neq 0 \end{cases} \quad (4)$$

If the system responses to a pulse sequence applied at each input are assembled to form the unit pulse response sequence denoted by

$$Y(0), Y(1), Y(2), \dots$$

then the input-output relationship can be expressed in terms of this pulse response sequence. Assuming zero initial conditions, this is given by

$$\begin{aligned} y(k) &= Y(0)u(k) + Y(1)u(k-1) + Y(2)u(k-2) + \dots \\ &= \sum_{i=0}^{\infty} Y(i)u(k-i) \end{aligned} \quad (5)$$

It can be seen from Eq. (5) that the contribution to the output at time step k by the input applied at time step k and at previous time steps $k-1, k-2, \dots$ are weighted by the pulse response sequence. For this reason the pulse response sequence is also known as the weighting sequence, and this input-output description is called the weighting sequence description. If the system is asymptotically stable, then the infinite summation in Eq. (5) can be approximated by a finite one since the weighting sequence may be truncated after a finite number of time steps. For lightly damped systems, however, the retained portion of the weighting sequence can be excessively large, and this description becomes cumbersome. Taking the Z-transform of both sides of Eq. (5) assuming zero initial conditions yields

$$\begin{aligned} \mathbf{Y}(z) &= [Y(0) + Y(1)z^{-1} + Y(2)z^{-2} + \dots] \mathbf{U}(z) \\ &= \mathbf{H}(z) \mathbf{U}(z) \end{aligned} \quad (6)$$

Thus the Z-transform of the weighting sequence is precisely the transfer function of the system in the z-domain. Furthermore, it is possible to relate the weighting sequence of a system to the coefficients of its difference representation in the time domain. Using Eq. (6) and Eq. (3) yields

$$(I + A_1 z^{-1} + \dots + A_n z^{-n})(Y(0) + Y(1)z^{-1} + Y(2)z^{-2} + \dots) = B_0 + B_1 z^{-1} + \dots + B_n z^{-n}$$

By comparing like powers of z^{-1} , the following relationship between the weighting sequence and the coefficients of the difference equation representation in Eq. (1) is obtained.

$$\sum_{k=0}^i A_k Y(i-k) = \begin{cases} B_i, & i = 0, 1, 2, \dots, n \\ 0, & i > n \end{cases} \quad (7)$$

$$A_0 = I$$

The quantities $Y(k)$, $k = 0, 1, 2, \dots$ are called the *Markov parameters* of the system. It immediately follows from the above discussion that (1) the Markov parameters are precisely the values of the weighting sequence at different time steps, (2) the sequence of Markov parameters is the pulse response of the system to a unit pulse, (3) the Z-transform of the Markov parameter sequence is the system transfer function in the complex z-domain.

2.3 State-Space Representation

The input-output relationship can also be expressed in terms of a set of n simultaneous first order difference equations of the form

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \quad (8)$$

where the dimensions of A, B, C , and D are $n \times n$, $n \times m$, $q \times n$, and $q \times m$, respectively. Solving for the output $y(k)$ in terms of the previous inputs yields

$$y(k) = \sum_{i=1}^k CA^{i-1} Bu(k-i) + Du(k) \quad (9)$$

The Markov parameter sequence is simply the pulse response of the system, therefore, they must be unique for a given system. Comparing Eq. (9) to Eq. (5) and noting that $u(k) = 0$ for $k < 0$ yields

$$\begin{aligned} Y(0) &= D \\ Y(k) &= CA^{k-1}B, \quad k = 1, 2, 3, \dots \end{aligned} \quad (10)$$

Equations (10) express the Markov parameters in terms of the system discrete state space matrices (A, B, C, D) . The state space matrices, however, are not unique since the state vector is coordinate-dependent. For example, let the state vector be transformed by a coordinate transformation T

$$z(k) = Tx(k)$$

Then Eq. (8) can be written as

$$\begin{aligned} z(k+1) &= TAT^{-1}z(k) + TBu(k) \\ y(k) &= CT^{-1}z(k) + Du(k) \end{aligned}$$

which is a state space description $(TAT^{-1}, TB, CT^{-1}, D)$ relating $u(k)$ and $y(k)$ by a new state vector $z(k)$. The system Markov parameters computed using these state space equations are the same as before,

$$Y(k) = CT^{-1}(TAT^{-1})^{k-1}TB = CA^{k-1}B$$

Thus, there are an infinite number of state-space representations that produce the same input-output mapping given in Eq. (9). Of particular interest are the canonical forms where the Markov parameters appear explicitly in these representations. For example, consider the single variable case, where the transfer function $H(z)$ in Eq. (3) is rewritten as

$$H(z) = \frac{\beta_1 z^{n-1} + \beta_2 z^{n-2} + \dots + \beta_{n-1} z + \beta_n}{z^n + \alpha_1 z^{n-1} + \dots + \alpha_{n-1} z + \alpha_n} + d \quad (11)$$

The observability canonical form of the above system is

$$A_0 = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ -\alpha_n & -\alpha_{n-1} & -\alpha_{n-2} & \dots & -\alpha_1 \end{bmatrix}, \quad B_0 = \begin{bmatrix} Y(1) \\ Y(2) \\ Y(3) \\ \vdots \\ Y(n) \end{bmatrix} \quad (12)$$

$$C_0 = [1 \ 0 \ 0 \ \dots \ 0] \quad D_0 = d$$

where the coefficients in the input matrix are precisely the first n Markov parameters obtained from

$$\frac{\beta_1 z^{n-1} + \beta_2 z^{n-2} + \dots + \beta_{n-1} z + \beta_n}{z^n + \alpha_1 z^{n-1} + \dots + \alpha_{n-1} z + \alpha_n} = \sum_{i=1}^{\infty} Y(i) z^{-i} \quad (13)$$

However, in the multiple variable case, such a canonical representation with the Markov parameters appearing explicitly is not of minimum order.

It is also possible to see the Markov parameters in the transfer function description of the system in state space format as given in Eq. (8). Taking the Z-transform of Eq. (8) yields $Y(z) = H(z)U(z)$ where the transfer function $H(z)$ can be expressed as

$$\begin{aligned} H(z) &= D + C(zI - A)^{-1} B \\ &= D + CBz^{-1} + CABz^{-2} + CA^2Bz^{-3} + \dots \end{aligned} \quad (14)$$

Finally, consider the case where the system is asymptotically stable. In this case,

$$\lim_{k \rightarrow \infty} CA^k B = 0$$

The input-output description in Eq. (9) can be approximated by a finite number of Markov parameters

$$\begin{aligned} y(k) &= Du(k) + CBu(k-1) + CABu(k-2) + \dots + CA^{p-1}Bu(k-p) \\ &= \sum_{i=1}^p CA^{i-1}Bu(k-i) + Du(k) \end{aligned} \quad (15)$$

where p is sufficiently large such that $CA^k B \approx 0$, $k \geq p$. By comparing with Eq. (1), this description is a special case of the linear difference description with

$$\begin{aligned} A_i &= 0, \quad B_0 = D \\ B_i &= Y(i) = CA^{i-1}B, \quad i = 1, 2, \dots, p \end{aligned} \quad (16)$$

The B_i coefficients appear explicitly as the system Markov parameters. Note that this description with a finite number of system Markov parameters is possible only for asymptotically stable systems.

2.4 Observer Equation for Input-Output Representation

The system input-output relation can also be described by a set of state space equations derived as follows. Add and subtract the term $My(i)$ to the right hand side of the state equation in Eq. (8) to yield

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + My(k) - My(k) \\ &= (A + MC)x(k) + (B + MD)u(k) - My(k) \end{aligned} \quad (17)$$

Define the following matrices

$$\bar{A} = A + MC \quad , \quad \bar{B} = [B + MD, -M] \quad , \quad v(k) = \begin{bmatrix} u(k) \\ y(k) \end{bmatrix} \quad (18)$$

Then the original system becomes

$$\begin{aligned} x(k+1) &= \bar{A}x(k) + \bar{B}v(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \quad (19)$$

If M is a matrix such that $\bar{A} = A + MC$ is asymptotically stable then the input-output relation can be approximated by a finite number of parameters $C\bar{A}^{i-1}\bar{B}$, say p of them.

$$y(k) = \sum_{i=1}^p C\bar{A}^{i-1}\bar{B}v(k-i) + Du(k) \quad (20)$$

Noting that $v(k)$ contains both $u(k)$ and $y(k)$, Eq. (20) can be rewritten as

$$y(k) + \sum_{i=1}^p C(A + MC)^{i-1}My(k-i) = \sum_{i=1}^p C(A + MC)^{i-1}(B + MD)u(k-i) + Du(k) \quad (21)$$

Observe that Eq. (20) is in a linear difference form as in Eq. (1) with the coefficients given by

$$\begin{aligned} A_i &= C(A + MC)^{i-1}M \quad , \quad B_0 = D \\ B_i &= C(A + MC)^{i-1}(B + MD) \quad , \quad i = 1, 2, \dots, p \end{aligned} \quad (22)$$

The role of the matrix M in the above development can be interpreted in terms of an observer. Consider the system given in Eq. (8). It has an observer of the form

$$\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + Bu(k) - M[y(k) - \hat{y}(k)] \\ \hat{y}(k) &= C\hat{x}(k) + Du(k) \end{aligned} \quad (23)$$

The state equation in Eq. (23) can be expressed as

$$\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + Bu(k) - MC[x(k) - \hat{x}(k)] \\ &= (A + MC)\hat{x}(k) + Bu(k) - M[y(k) - Du(k)] \\ &= (A + MC)\hat{x}(k) + (B + MD)u(k) - My(k) \end{aligned} \quad (24)$$

Defining the state estimation error $e(k) = x(k) - \hat{x}(k)$, the equation that governs $e(k)$ is

$$\begin{aligned} e(k+1) &= Ax(k) + Bu(k) - [(A + MC)\hat{x}(k) + (B + MD)u(k) - My(k)] \\ &= (A + MC)e(k) \end{aligned}$$

If the system (8) is observable, then the matrix M may be chosen to place the eigenvalues of $A + MC$ in any desired (symmetric) configuration. If the matrix M is such that $A + MC$ is asymptotically stable, then the estimated state $\hat{x}(k)$ tends to the true state $x(k)$ as k tends to infinity. Equation (24) then becomes

$$x(k+1) = (A + MC)x(k) + (B + MD)u(k) - My(k)$$

which is exactly the same as Eq. (17). From this analysis, the matrix M can be interpreted as an observer gain.

The parameters $C\bar{A}^{i-1}\bar{B}$ in Eq. (19) are the Markov parameters of an observer system, hence they are referred to as *observer Markov parameters*. They are denoted by

$$\bar{Y}(0) = D, \quad \bar{Y}(k) = C\bar{A}^{k-1}\bar{B}, \quad k = 1, 2, \dots, p \quad (25)$$

The extra freedom inherent in these parameters can be exploited to develop various identification algorithms.⁵⁻¹¹ Consider the case where M corresponds to a deadbeat observer gain, the observer Markov parameters become identically zero after a finite number of terms. For lightly damped structures, this means that the system can be described by a reduced number of observer Markov parameters $\bar{Y}(k)$ instead of an otherwise large number of the usual system Markov parameters $Y(k)$. For this reason, the observer Markov parameters are useful in the development of various system identification algorithms. The relationship between the Markov parameters of the observer system and those of the actual system will be presented in the following section.

3. Relationship between the Markov Parameters of the Observer and Actual System

The observer Markov parameters contain information about the system and the observer. It is possible to recover both the system and the observer gain from a given set of observer Markov parameters. This section describes this relationship.

3.1 Recovery of System Markov Parameters from Observer Markov Parameters

In this section, the basic relationship between the Markov parameters of the observer and actual system is described. First, it is shown in the following that given a set of observer Markov parameters, the system Markov parameters can be recovered uniquely. The direct transmission term D is simply

$$D = Y(0) = \bar{Y}(0) \quad (26)$$

For ease of presentation, the observer Markov parameters are partitioned as

$$\begin{aligned} \bar{Y}(k) &= C\bar{A}^{k-1}\bar{B} \\ &= [C(A + MC)^{k-1}(B + MD) \quad -C(A + MC)^{k-1}M] \\ &= [\bar{Y}^{(1)}(k) \quad \bar{Y}^{(2)}(k)] \end{aligned} \quad (27)$$

The Markov parameter $Y(1) = CB$ of the system is simply

$$Y(1) = CB = \bar{Y}^{(1)}(1) + \bar{Y}^{(2)}(1)Y(0)$$

To obtain the next Markov parameter $Y(2) = CAB$, consider $\bar{Y}^{(2)}$

$$Y(2) = CAB = \bar{Y}^{(1)}(2) + \bar{Y}^{(2)}(1)Y(1) + \bar{Y}^{(2)}(2)Y(0)$$

Similarly,

$$\begin{aligned} Y(3) &= CA^2B \\ &= \bar{Y}^{(1)}(3) - CM CAB - C(A + MC)MCB - C(A + MC)^2 MD \\ &= \bar{Y}^{(1)}(3) + \bar{Y}^{(2)}(1)Y(2) + \bar{Y}^{(2)}(2)Y(1) + \bar{Y}^{(2)}(3)Y(0) \end{aligned}$$

By induction, the general relationship between the actual system Markov parameters and the observer Markov parameters can be shown to be

$$Y(k) = \bar{Y}^{(1)}(k) + \sum_{i=1}^k \bar{Y}^{(2)}(i)Y(k-i) \quad (28)$$

where $\bar{Y}^{(k)}$ are considered to be zero for $k > p$. The above recursive equation can be written in matrix form as

$$\begin{bmatrix} I & & & & \\ -\bar{Y}^{(2)}(1) & I & & & \\ -\bar{Y}^{(2)}(2) & -\bar{Y}^{(2)}(1) & \ddots & & \\ \vdots & \vdots & \ddots & I & \\ -\bar{Y}^{(2)}(k-1) & -\bar{Y}^{(2)}(k-2) & \dots & -\bar{Y}^{(2)}(1) & I \end{bmatrix} \begin{bmatrix} Y(1) \\ Y(2) \\ Y(3) \\ \vdots \\ Y(k) \end{bmatrix} = \begin{bmatrix} \bar{Y}^{(1)}(1) \\ \bar{Y}^{(1)}(2) \\ \bar{Y}^{(1)}(3) \\ \vdots \\ \bar{Y}^{(1)}(k) \end{bmatrix} + \begin{bmatrix} \bar{Y}^{(2)}(1) \\ \bar{Y}^{(2)}(2) \\ \bar{Y}^{(2)}(3) \\ \vdots \\ \bar{Y}^{(2)}(k) \end{bmatrix} Y(0) \quad (29)$$

The left most matrix in the above equation is square and full rank, which implies that from a given set of observer Markov parameters $\bar{Y}^{(k)}$, the system Markov parameters $Y(k)$ can be uniquely recovered.

Furthermore, let the matrices $\bar{\mathbf{Y}}^{(2)}$, and \mathbf{Y} be defined as follows

$$\begin{aligned} \bar{\mathbf{Y}}^{(2)} &= [-\bar{Y}^{(2)}(p) \quad -\bar{Y}^{(2)}(p-1) \quad -\bar{Y}^{(2)}(p-2) \quad \dots \quad -\bar{Y}^{(2)}(1)] \\ \mathbf{Y} &= [Y(p+2) \quad Y(p+3) \quad Y(p+4) \quad \dots \quad Y(2p+1)] \end{aligned} \quad (30)$$

then from Eq. (28)

$$\bar{\mathbf{Y}}^{(2)}\mathbf{H}(1) = \mathbf{Y} \quad (31)$$

where

$$\mathbf{H}(1) = \begin{bmatrix} Y(2) & Y(3) & \cdots & Y(p+1) \\ Y(3) & Y(4) & \cdots & Y(p+2) \\ \vdots & \vdots & \ddots & \vdots \\ Y(p+1) & Y(p+2) & \cdots & Y(2p) \end{bmatrix} \quad (32)$$

Examining the rank properties of Eq. (31) reveals that to describe a system of order n , the number of observer Markov parameters p must be such that $qp \geq n$ where q is the number of outputs. Furthermore, the maximum order of a system that can be described with p observer Markov parameters is qp . The implication of this analysis is that for multiple-output systems, the number of observer Markov parameters required to be identified can be substantially less than the true order of the system. Specifically, the minimum number of observer Markov parameters that can describe the system is p_{\min} , which is the smallest value of p such that $qp_{\min} \geq n$.

3.2 Recovery of Observer Gain from Observer Markov Parameters

Since the observer Markov parameters contain information of both the system and observer gain, it is possible also to recover the observer gain M from the observer Markov parameters. First, the sequence

$$Y_M(k) = CA^{k-1}M, \quad k = 1, 2, \dots \quad (33)$$

can be computed from the observer Markov parameter sequence $\bar{Y}(k) = C\bar{A}^{k-1}\bar{B}$, $k = 1, 2, \dots$ as follows

$$Y_M(k) = -\bar{Y}^{(2)}(k) + \sum_{i=1}^{k-1} \bar{Y}^{(2)}(i)Y_M(k-i) \quad (34)$$

The parameters in Eq. (33) contain information about the observer gain, therefore they are referred to as *observer gain Markov parameters*. Given A , C , and a sufficient number of observer gain Markov parameters $Y_M(k)$ computed from Eq. (34), the observer gain M can be obtained from

$$M = (V_k^T V_k)^{-1} V_k^T Y_M \quad (35)$$

where

$$V_k = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^k \end{bmatrix} \quad Y_M = \begin{bmatrix} CM \\ CAM \\ \vdots \\ CA^k M \end{bmatrix}$$

provided V_k is of rank n . Alternatively, the observer gain matrix M can be realized simultaneously with the system matrices A , B , C from the combined Markov parameters

$$\begin{aligned}
P(k) &= [CA^{k-1}B \quad CA^{k-1}M] \\
&= CA^{k-1}[B \quad M] , \quad k = 1, 2, \dots
\end{aligned} \tag{36}$$

This problem will be revisited in a later section.

4. Relationship between Discrete and Continuous-Time Models

The results presented so far are concerned with discrete-time models. However, all structural systems are continuous, but information available for identification is in the form of sampled data. In this section, the relationship between continuous and discrete-time models is described.

4.1 Continuous-Time Dynamic Equations

Consider the equations of motion for a finite-dimensional linear dynamic system of the form

$$M\ddot{w}(t) + D\dot{w}(t) + Kw(t) = B_f u(t) \tag{37}$$

where M , D , and K denote the system mass, damping, and stiffness matrices, respectively. If the dynamic system is measured by output quantities such as position, velocity, and acceleration then in general the set of observation variables can be written in matrix form as

$$\begin{aligned}
y(t) &= C_1 w(t) + C_2 \dot{w}(t) + C_3 \ddot{w}(t) \\
&= [C_1 - C_3 M^{-1}K]w(t) + [C_2 - C_3 M^{-1}D]\dot{w}(t) + C_3 M^{-1}B_f u(t)
\end{aligned}$$

Equation (37) together with the output equation can be written in first order form as

$$\begin{aligned}
\dot{x}(t) &= A_c x(t) + B_c u(t) \\
y(t) &= Cx(t) + Du(t)
\end{aligned} \tag{38}$$

where

$$\begin{aligned}
x(t) &= \begin{bmatrix} w(t) \\ \dot{w}(t) \end{bmatrix} , \quad A_c = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}D \end{bmatrix} , \quad B_c = \begin{bmatrix} 0 \\ M^{-1}B_f \end{bmatrix} \\
C &= [C_1 - C_3 M^{-1}K \quad C_2 - C_3 M^{-1}D] , \quad D = C_3 M^{-1}B_f
\end{aligned} \tag{39}$$

This is the continuous-time equations in state space format. The continuous-time state space equation can be converted to the discrete-time equation by the following procedure.

4.2 Conversion of Continuous to Discrete-Time State Space Model

The general solution for $x(t)$ at any time t for any input $u(t)$ from any initial condition $x(t_0)$ is

$$x(t) = e^{A_c(t-t_0)}x(t_0) + \int_{t_0}^t e^{A_c(t-\tau)}B_c u(\tau)d\tau$$

Consider discrete sampling time intervals $0, T, 2T, \dots$. To see how $x(t)$ changes from one time step to the next, replace $t_0 = kT$ and $t = (k+1)T$ in the above equation. Furthermore, assume a zero-order hold on the input, i.e., the input $u(kT)$ is held constant over the time interval from kT to $(k+1)T$, one obtains

$$\begin{aligned} x[(k+1)T] &= e^{A_c T} x(kT) + \int_{kT}^{(k+1)T} e^{A_c[(k+1)T-\tau]} B_c u(\tau) d\tau \\ &= e^{A_c T} x(kT) + \left(\int_0^T e^{A_c \tau} d\tau B_c \right) u(kT) \end{aligned}$$

Using a simplified notation k for the time argument kT ,

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \quad (40)$$

where the output equation is written at every time step, and

$$A = e^{A_c T}, \quad B = \int_0^T e^{A_c \tau} d\tau B_c \quad (41)$$

and the C and D matrices are the same as in the continuous-time representation, Eq. (38). Equation (40) is a discrete-time state space equation with a sampling interval T for the continuous-time dynamic system in Eq. (37).

In practice, to compute the matrix exponential, one can use an approximation of the form¹⁹

$$e^{A_c T} = \left[R_{\vartheta\vartheta} \left(\frac{A_c T}{2^j} \right) \right]^{2^j}, \quad \vartheta, j \geq 0 \quad (42)$$

where $R_{\vartheta\vartheta}(z)$ is the (ϑ, ϑ) Padé approximant to e^z

$$R_{\vartheta\vartheta}(z) = \frac{\sum_{k=0}^{\vartheta} c_k z^k}{\sum_{k=0}^{\vartheta} c_k (-z)^k}, \quad c_k = \frac{(2\vartheta - k)! \vartheta!}{(2\vartheta)! k! (\vartheta - k)!} \quad (43)$$

In fact, the discrete-time A and B can be computed from the continuous-time A_c and B_c in one step by making use of the following identity

$$\Phi(t=T) = e^{\begin{bmatrix} A_c & B_c \\ 0 & 0 \end{bmatrix} T} = \begin{bmatrix} A & B \\ 0 & I \end{bmatrix} \quad (44)$$

The above identity can be proved by first noting that $\Phi(t)$ is the solution of the following matrix differential equation

$$\frac{d\Phi}{dt} = \begin{bmatrix} A_c & B_c \\ 0 & 0 \end{bmatrix} \Phi \quad \Phi(0) = I \quad (45)$$

Denote Φ_{11} , Φ_{12} , Φ_{21} , Φ_{22} as the partitions of Φ ,

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}$$

From Eq. (44)

$$\frac{d\Phi_{11}}{dt} = A_c \Phi_{11}, \quad \frac{d\Phi_{12}}{dt} = A_c \Phi_{12} + B_c \Phi_{22}, \quad \frac{d\Phi_{21}}{dt} = 0, \quad \frac{d\Phi_{22}}{dt} = 0 \quad (46)$$

Application of the initial conditions $\Phi_{11} = I$, $\Phi_{21} = 0$, $\Phi_{22} = I$ yields $\Phi_{11} = A$, $\Phi_{21} = 0$, $\Phi_{22} = I$. To show $\Phi_{12} = B$, pre-multiply both sides of the Φ_{12} equation by $e^{-A_c t}$ and rearrange terms

$$e^{-A_c t} \frac{d\Phi_{12}}{dt} - e^{-A_c t} A_c \Phi_{12} = e^{-A_c t} B_c$$

Integrating both sides of the above equation from $t = 0$ to $t = T$, recognizing that the left hand side is a perfect differential, yields

$$\int_0^T \frac{d}{dt} [e^{-A_c t} \Phi_{12}(t)] dt = e^{-A_c t} \Phi_{12}(t) \Big|_{t=0}^{t=T} = \int_0^T e^{-A_c t} B_c dt$$

Solve for $\Phi_{12}(T)$, noting that $\Phi_{12}(0) = 0$,

$$\Phi_{12}(T) = \int_0^T e^{A_c(T-t)} B_c dt = \int_0^T e^{A_c \tau} B_c d\tau = B \quad (47)$$

The last step in Eq. (47) is obtained by a change of integration variable from $(T - t)$ to τ .

4.3 Conversion of Discrete to Continuous-Time State Space Model

A realized discrete-time system matrix A can be converted to a continuous-time one. Assume that A can be diagonalized by a matrix P such that $P^{-1}AP = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, then

$$\Lambda = P^{-1} e^{A_c T} P = e^{P^{-1} A_c P T} = e^{\text{diag}\left(\frac{\ln \lambda_1}{T}, \frac{\ln \lambda_2}{T}, \dots, \frac{\ln \lambda_n}{T}\right) T}$$

A continuous-time system matrix A_c can be obtained from

$$A_c = P \text{diag} \left(\frac{\ln \lambda_1}{T}, \frac{\ln \lambda_2}{T}, \dots, \frac{\ln \lambda_n}{T} \right) P^{-1} \quad (48)$$

The eigenvalues of the continuous-time A_c are

$$\begin{aligned} \gamma_j &= \sigma_j \pm i\omega_j = \frac{\ln \lambda_i}{T} \\ &= \text{Re} \left(\frac{\ln \lambda_i}{T} \right) \pm i \left\{ \text{Im} \left(\frac{\ln \lambda_i}{T} \right) \pm \frac{2\ell\pi}{T} \right\}, \quad \ell = 0, 1, 2, \dots \end{aligned} \quad (49)$$

There is no unique continuous-time system that produces a given discrete-time model. Under the assumption that the sampling rate is sufficiently fast such that all frequencies in the signals being identified fall below the Nyquist frequency, an appropriate value of ℓ may be chosen. In practice, there will always be frequencies outside the Nyquist frequency, and this choice of ℓ will misinterpret as frequencies within the Nyquist frequency and thus produce modelling errors. This phenomenon is known as aliasing. In the identification of structures, the complex parts of the eigenvalues are the modal frequencies, and the real parts are the modal damping factors characterizing the decay rates of the modal oscillations.

To convert the discrete B to a continuous B_c , first write $A = e^{A_c T}$ as a convergent power series

$$A = e^{A_c T} = I + A_c T + \frac{A_c^2 T^2}{2!} + \frac{A_c^3 T^3}{3!} + \dots \quad (50)$$

Therefore,

$$A - I = A_c T + \frac{A_c^2 T^2}{2!} + \frac{A_c^3 T^3}{3!} + \dots$$

Hence

$$(A - I)B_c = A_c B_c T + \frac{A_c^2 B_c T^2}{2!} + \frac{A_c^3 B_c T^3}{3!} + \dots \quad (51)$$

Integrate the second equation in Eq. (41) term by term, and pre-multiply by A_c

$$A_c B = A_c B_c T + \frac{A_c^2 B_c T^2}{2} + \frac{A_c^3 B_c T^3}{3 \times 2!} + \dots \quad (52)$$

Comparing Eq. (51) and Eq. (52) immediately yields

$$(A - I)B_c = A_c B$$

The continuous-time input matrix B_c is

$$B_c = (A - I)^{-1} A_c B \quad (53)$$

In this section, to derive Eq. (48) and Eq. (53), the matrix A is required to be diagonalizable. This requirement excludes the presence of rigid body modes in the model. To account for rigid body modes, a different procedure must be used to convert from a discrete to a continuous-time presentation. Note that the reverse procedure to convert from the continuous to a discrete-time presentation formulated in the previous section using matrix exponential does not have this limitation.

5. Toeplitz and Hankel Matrices

Two important matrices often used in system identification and control are the Toeplitz matrix and the Hankel matrix. They are simply matrices of Markov parameters of particular structures.

5.1 The Toeplitz Matrix

Often associated with the Markov parameters is a lower block triangular Toeplitz matrix. The input and output histories of the discrete system given Eq. (8) satisfy the relation

$$\begin{bmatrix} y(k) \\ y(k+1) \\ \vdots \\ y(k+n-1) \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} x(k) + \begin{bmatrix} D & & & \\ CB & D & & \\ \vdots & \ddots & \ddots & \\ CA^{n-2}B & \cdots & CB & D \end{bmatrix} \begin{bmatrix} u(k) \\ u(k+1) \\ \vdots \\ u(k+n-1) \end{bmatrix} \quad (54)$$

The observability matrix V_{n-1} and the Toeplitz matrix $T(n-1)$ formed by the Markov parameters are defined as

$$V_{n-1} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \quad T(n-1) = \begin{bmatrix} Y(0) & & & \\ Y(1) & Y(0) & & \\ \vdots & \ddots & \ddots & \\ Y(n-1) & \cdots & Y(1) & Y(0) \end{bmatrix} \quad (55)$$

with the blank spaces denoting zero elements. The Toeplitz matrix is also known as the pulse response matrix.

5.2 The Hankel Matrix

Another important matrix of Markov parameters which is often encountered in realization theory is the Hankel matrix. The Hankel matrix is defined as

$$H(k-1) = \begin{bmatrix} Y(k) & Y(k+1) & \cdots & Y(k+s) \\ Y(k+1) & Y(k+2) & \cdots & Y(k+s+1) \\ \vdots & \vdots & & \vdots \\ Y(k+r) & Y(k+r+1) & \cdots & Y(k+r+s) \end{bmatrix} \quad (56)$$

If $r+1 \geq n$ and $s+1 \geq n$, the Hankel matrix is of rank n . The Hankel matrix can be factorized in terms of the system observability matrix V_r and controllability matrix W_s as

$$\mathbf{H}(k) = V_r A^k W_s \quad (57)$$

where

$$V_r = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^r \end{bmatrix} \quad W_s = [B \quad AB \quad \cdots \quad A^{s-1}B]$$

For the scalar case considered in the previous section, it is interesting to point out a relationship between the system parameters, the Hankel matrix, and the Toeplitz matrix for the system characterized by the transfer function given in Eq. (11). They are related by

$$\begin{aligned} \mathbf{H}(0) [\alpha_n \quad \alpha_{n-1} \quad \cdots \quad \alpha_1]^T &= [Y(n+1) \quad Y(n+2) \quad \cdots \quad Y(2n)]^T \\ [\beta_1 \quad \beta_2 \quad \cdots \quad \beta_n]^T &= \bar{\mathbf{T}}(n) [1 \quad \alpha_1 \quad \cdots \quad \alpha_{n-1}]^T \end{aligned} \quad (58)$$

where

$$\mathbf{H}(0) = \begin{bmatrix} Y(1) & Y(2) & \cdots & Y(n) \\ Y(2) & Y(3) & \cdots & Y(n+1) \\ \vdots & \vdots & \ddots & \vdots \\ Y(n) & Y(n+1) & \cdots & Y(2n-1) \end{bmatrix} \quad \bar{\mathbf{T}}(n) = \begin{bmatrix} Y(1) & & & \\ Y(2) & Y(1) & & \\ \vdots & \ddots & \ddots & \\ Y(n) & \cdots & Y(2) & Y(1) \end{bmatrix}$$

6. Realization from Markov Parameters

The problem of finding a state space model from a given sequence of Markov parameters is known as realization. The theory of minimal realization finds such a model with minimum dimensions. A realization technique applicable to identification of structures is the Eigensystem Realization Algorithm (ERA) which is summarized in the following.

6.1 State Space System Realization by ERA

ERA considers the problem where a sequence of Markov parameters $Y(1), Y(2), Y(3), \dots$ is given. The objective is to find a state-space realization A, B, C such that the relation given in Eq. (10) holds. The algorithm begins with a Hankel matrix defined as in Eq. (56). The order of the system n is determined from the singular value decomposition of $\mathbf{H}(0)$

$$\mathbf{H}(0) = U \Sigma V^T = U_1 \Sigma_1 V_1^T \quad (59)$$

The columns of U_1 and V_1 in Eq. (59) are orthonormal, and correspond to the positive singular values in an $n \times n$ diagonal matrix Σ_1 . A discrete-time minimal order realization of the system can be shown to be

$$A_r = \Sigma_1^{-1/2} U_1^T H(1) V_1 \Sigma_1^{-1/2}, \quad B_r = \Sigma_1^{1/2} V_1^T E_m, \quad C_r = E_q^T U_1 \Sigma_1^{1/2} \quad (60)$$

where $E_\alpha^T = [I_\alpha \ O_\alpha \ \dots \ O_\alpha]$, $\alpha = m, q$. Thus, for a finite dimensional system, knowledge of a sufficient number of Markov parameters is adequate to obtain a state-space realization for the system.

An important aspect in connection with system realization from Markov parameters is the order of the minimal realization. In the noise-free case, given a sufficient number of Markov parameters, the minimal order n is equal to the rank of the Hankel matrix, which is equal to the number of positive singular values in Σ_1 . Equation (59) holds exactly. This is evident from the factorization given in Eq. (57) and the fact that a minimal order realization must be observable and controllable. In practice, however, all structural systems have noises and nonlinearities, the problem of rank determination is not trivial. The singular value decomposition step in the above realization procedure is normally used to determine this order. Certain smaller singular values corresponding to less significant modes are attributed to noises, and truncated. The number of retained singular values then determine the system order. The corresponding retained columns of U and V in the singular value decomposition of $H(0)$ are used in Eq. (60) to obtain a realization.

6.2 State Space System/Observer Realization

Realization can be thought of as a factorization of a parameter sequence of the form $Y(k) = CA^{k-1}B$, to obtain a set of (C, A, B) that preserves the prescribed relationship between the parameters in the sequence. Therefore, the theory is applicable to factorization of all such sequences whose parameters obey the same prescribed relationship. The ERA procedure can be applied to the combined Markov parameter sequence $P(k)$ given in Eq. (36) by first forming the Hankel matrix of $P(k)$

$$H_c(k-1) = \begin{bmatrix} P(k) & P(k+1) & \dots & P(k+s) \\ P(k+1) & P(k+2) & \dots & P(k+s+1) \\ \vdots & \vdots & & \vdots \\ P(k+r) & P(k+r+1) & \dots & P(k+r+s) \end{bmatrix} \quad (61)$$

The following realization will simultaneously identify A, B, C and the observer gain M

$$A_r = D_1^{-1/2} P_1^T H_c(1) Q_1 D_1^{-1/2}, \quad [B_r \ M_r] = D_1^{1/2} Q_1^T E_{m+q}, \quad C_r = E_q^T P_1 D_1^{1/2} \quad (62)$$

where $H_c(0) = PDQ^T = P_1 D_1 Q_1^T$, and $E_\alpha^T = [I_\alpha \ O_\alpha \ \dots \ O_\alpha]$, $\alpha = m+q, q$.

In the previous development the set of parameters is precisely the Markov parameter sequence of a discrete system. In system identification, the Markov parameters are obtained or computed from measurement data. The determination of the Markov parameters from sampled data is discussed in the following section.

7. Determination of Markov Parameters from Sampled Data

Under certain conditions, it is possible to interpret the sampled data directly as Markov parameters. In some other cases, the Markov parameters can be computed from input-output data. In particular, the following cases are considered.

7.1 Impulse Test Input

Let the system given in Eq. (38) be subjected to an impulse in the j -th input, starting from zero initial conditions. The response is recorded at evenly spaced discrete-time intervals, and denoted by $y^{(j)}(kT)$, $k = 0, 1, 2, \dots$ beginning with $y^{(j)}(0) = y^{(j)}(0^+)$ immediately after the impulse has entered the system. The sampled response for the j -th input impulse is

$$\begin{aligned} y^{(j)}(kT) &= Ce^{A_k T} x(0) + Ce^{A_k T} \int_0^{kT} e^{-A_c \tau} B_c^{(j)} u(\tau) d\tau + Du(kT) \\ &= C(e^{A_k T})^k B_c^{(j)} = CA^k B_c^{(j)} \end{aligned} \quad (63)$$

where $B_c^{(j)}$ denotes the j -th column of the input matrix B_c , $j = 1, 2, \dots, m$. The same impulse test is applied to every input in the system, one at a time. The data is then assembled to form the following data sequence, which has the same structure as that of the Markov parameters

$$\begin{aligned} [y^{(1)}(kT) \quad y^{(2)}(kT) \quad \dots \quad y^{(m)}(kT)] &= [CA^k B_c^{(1)} \quad CA^k B_c^{(2)} \quad \dots \quad CA^k B_c^{(m)}] \\ &= CA^k [B_c^{(1)} \quad B_c^{(2)} \quad \dots \quad B_c^{(m)}] \\ &= CA^k B_c, \quad k = 0, 1, 2, 3, \dots \end{aligned} \quad (64)$$

where A is the discrete-time system matrix for the system in Eq. (38) corresponding to the sampling period T , $A = e^{A_c T}$, but B_c is the continuous-time input matrix, and C the output matrix which is the same for both discrete and continuous-time cases. Note that the data is recorded after the impulse has entered the system, therefore, information regarding D is not available. Since a true impulse is actually not available in practice, for a sufficiently rich input, the needed data sequence can be obtained by taking the inverse fast Fourier transform (FFT) of the ratio of the FFT transforms of the input and output measurements.

7.2 Unit Pulse Test Input

A way to obtain the true Markov parameter sequence is by applying a unit pulse for the j -th input

$$u_j(kT) = \begin{cases} 1 & , k = 0 \\ 0 & , k = 1, 2, 3, \dots \end{cases} \quad (65)$$

with a zero-order hold on the input to maintain the input values during each sampling interval. The discrete-time equivalence of the continuous-time system of Eq. (38) for the j -th input is

$$\begin{aligned}x(k+1) &= Ax(k) + B^{(j)}u_j(k) \\y^{(j)}(k) &= Cx(k) + D^{(j)}u_j(k)\end{aligned}$$

where $B^{(j)}$ and $D^{(j)}$ denote the j -th column of B and D , respectively. The corresponding response is measured at discrete sampling intervals, and denoted by $y^{(j)}(kT)$, $k = 1, 2, 3, \dots$. With zero initial conditions, a separate test is conducted for each input. The sampled responses of the continuous-time system are then assembled to form the data sequence

$$[y^{(1)}(0) \ y^{(2)}(0) \ \dots \ y^{(m)}(0)] = [D^{(1)} \ D^{(2)} \ \dots \ D^{(m)}] = D \quad (66)$$

and

$$\begin{aligned}[y^{(1)}(kT) \ y^{(2)}(kT) \ \dots \ y^{(m)}(kT)] &= [CA^{k-1}B^{(1)} \ CA^{k-1}B^{(2)} \ \dots \ CA^{k-1}B^{(m)}] \\&= CA^{k-1}[B^{(1)} \ B^{(2)} \ \dots \ B^{(m)}] \\&= CA^{k-1}B, \quad k = 1, 2, 3, \dots\end{aligned} \quad (67)$$

Observe that this sequence is precisely the true Markov parameter sequence of the system discrete-time state space model with a sampling interval equal to the duration of the unit pulse test input.

7.3 Extended Unit Pulse Test Input

To put more energy into the input, consider a unit pulse in the j -th input held for p time steps

$$u_j(kT) = \begin{cases} 1 & , \ k = 0, 1, 2, \dots, p-1 \\ 0 & , \ k = p, p+1, p+2, \dots \end{cases} \quad (68)$$

With zero initial conditions, the response sequence at time steps $0, T, 2T, \dots, (p-1)T$ is

$$y^{(j)}(0) = D^{(j)}$$

and

$$\begin{aligned}y^{(j)}(T) &= CB^{(j)} + D^{(j)} \\y^{(j)}(2T) &= CAB^{(j)} + CB^{(j)} + D^{(j)} \\&\vdots \\y^{(j)}[(p-1)T] &= CA^{p-2}B^{(j)} + \dots + CB^{(j)} + D^{(j)}\end{aligned}$$

and for time steps beyond pT

$$\begin{aligned}y^{(j)}(pT) &= CA^{p-1}B^{(j)} + \dots + CB^{(j)} \\y^{(j)}[(p+1)T] &= CA^pB^{(j)} + \dots + CAB^{(j)} \\&\vdots\end{aligned}$$

The Markov parameter sequence associated with the j -th input is simply

$$\begin{aligned}
D^{(j)} &= y^{(j)}(0) \\
CB^{(j)} &= y^{(j)}(T) - y^{(j)}(0) \\
&\vdots \\
CA^{p-2}B^{(j)} &= y^{(j)}[(p-1)T] - y^{(j)}[(p-2)T]
\end{aligned} \tag{69}$$

and for $k = p-1, p, \dots$

$$CA^k B^{(j)} = y^{(j)}[(k+1)T] - \sum_{i=1}^{p-1} CA^{k-i} B^{(j)} \tag{70}$$

The results for each input are then assembled to obtain the system discrete-time Markov parameters.

7.4 Free Decay Response

Consider the free response of the system given in Eq. (40) from a non-zero initial condition, the sampled response is

$$y(kT) = CA^k x(0), \quad k = 0, 1, 2, \dots \tag{71}$$

The above sequence has the same mathematical structure as the Markov parameter sequence, except that the initial condition $x(0)$ plays the role of B . An identification procedure applied to the above sequence will produce a realization of C, A , and the initial condition $x(0)$.

7.5 Random Test Input

Consider the case where the system given in Eq. (40) is driven by an independent, white, zero-mean, stationary random input sequence with unit covariance

$$E\{u(kT)u(kT)^T\} = I, \quad k = 1, 2, 3, \dots \tag{72}$$

With a zero-order hold on the input, the resultant response is measured at discrete sampling intervals, $y(kT)$, $k = 1, 2, 3, \dots$. First, note that

$$E\{y(k)u(k)^T\} = E\{[Cx(k) + Du(k)]u(k)^T\} = D \tag{73}$$

Next, consider

$$E\{y(k+1)u(k)^T\} = E\{[CAx(k) + CBu(k) + Du(k+1)]u(k)^T\} = CB$$

In general, the Markov parameters of the system are related to the covariance between the input and output data by

$$E\{y(k+i)u(i)^T\} = CA^{k-1}B, \quad k = 1, 2, \dots \tag{74}$$

In practice, if the above relations are to be used to compute the Markov parameters, the expected values are replaced by sample covariance computation. This approximation is only valid when the input-output data sequences are stationary and the records are sufficiently long such that the ergodicity property of stationary random processes apply.

7.6 System Markov Parameters from General Input-Output Data

The problem of determining the Markov parameters from a general test input is can be addressed using various input-output models that are described in previous sections. Suppose that a set of $N+1$ measurements of $y(i)$ and $u(i)$, $i = 0, 1, \dots, N$, is available. Assuming zero initial conditions, the input-output data can be written in matrix form as

$$\mathbf{y} = \mathbf{Y}_{N-1} \mathbf{U}_N \quad (75)$$

where

$$\mathbf{y} = [y(0) \ y(1) \ \dots \ y(p) \ y(p+1) \ \dots \ y(N)]$$

$$\mathbf{Y}_{N-1} = [D \ CB \ \dots \ CA^{p-1}B \ CA^pB \ \dots \ CA^{N-1}B]$$

$$\mathbf{U}_N = \begin{bmatrix} u(0) & u(1) & \dots & u(p) & u(p+1) & \dots & u(N) \\ & u(0) & \dots & u(p-1) & u(p) & \dots & u(N-1) \\ & & \ddots & \vdots & \vdots & \ddots & \vdots \\ & & & u(0) & u(1) & \vdots & \vdots \\ & & & & u(0) & \ddots & \vdots \\ & & & & & \ddots & u(1) \\ & & & & & & u(0) \end{bmatrix}$$

In the single-input case, \mathbf{U}_N is a square matrix. The Markov parameters in \mathbf{Y}_{N-1} can be solved from this equation,

$$\mathbf{Y}_{N-1} = \mathbf{y} \mathbf{U}_N^{-1} \quad (76)$$

provided $u(0) \neq 0$ so that \mathbf{U}_N is full rank. This solution cannot be extended to the multiple-input problem since in this case \mathbf{Y}_{N-1} cannot be determined uniquely. If the system is asymptotically stable, and p is sufficiently large such that the Markov parameters $CA^{k-1}B$ can be neglected for $k > p$ then it is possible to solve for the Markov parameters $D, CB, CAB, \dots, CA^{p-1}B$ from the linear set of equations

$$\mathbf{y} = \mathbf{Y} \mathbf{U} \quad (77)$$

where

$$\mathbf{Y} = [D \ CB \ \dots \ CA^{p-1}B]$$

$$\mathbf{U} = \begin{bmatrix} u(0) & u(1) & \cdots & u(p) & u(p+1) & \cdots & u(N) \\ & u(0) & \cdots & u(p-1) & u(p-2) & \cdots & u(N-1) \\ & & \ddots & \vdots & \vdots & \vdots & \vdots \\ & & & u(0) & u(1) & \cdots & u(N-p) \end{bmatrix}$$

The solution for the system Markov parameters from Eq. (77) is

$$\mathbf{Y} = \mathbf{y}\mathbf{U}^T(\mathbf{U}\mathbf{U}^T)^{-1} \quad (78)$$

provided that the data record is sufficiently long, and the input is sufficiently rich such that the inverse in Eq. (78) exists. For lightly damped, large flexible structures with multiple sensors and actuators, this direct solution for the Markov parameters is not desirable in practice since the number system Markov parameters required in \mathbf{Y} may be excessively large.

If the initial conditions are not zero then a slightly different equation must be used in place of Eq. (78) to solve for the system Markov parameters. This is given as

$$\mathbf{Y} = \mathbf{y}_i \mathbf{U}_i^T (\mathbf{U}_i \mathbf{U}_i^T)^{-1} \quad (79)$$

where \mathbf{y}_i and \mathbf{U}_i are obtained by deleting the first p columns in \mathbf{y} and \mathbf{U} , respectively.

7.7 Observer Markov Parameters from General Input-Output Data (OKID)

The above mentioned problem can be overcome by solving for the observer Markov parameters instead. The observer equation counterpart of Eq. (77) is simply

$$\mathbf{y} = \bar{\mathbf{Y}}\mathbf{v} \quad (80)$$

where

$$\bar{\mathbf{Y}} = [D \quad C\bar{B} \quad \cdots \quad C\bar{A}^{p-1}\bar{B}]$$

$$\mathbf{v} = \begin{bmatrix} u(0) & u(1) & \cdots & u(p) & u(p+1) & \cdots & u(N) \\ v(0) & v(1) & \cdots & v(p-1) & v(p-2) & \cdots & v(N-1) \\ & \ddots & & \vdots & \vdots & \vdots & \vdots \\ & & & v(0) & v(1) & \cdots & v(N-p) \end{bmatrix}$$

where $v(k)$ is defined as in Eq. (18). A solution for the observer Markov parameters is

$$\bar{\mathbf{Y}} = \mathbf{y}\mathbf{v}^T(\mathbf{v}\mathbf{v}^T)^{-1} \quad (81)$$

If the initial conditions are not zero then a slightly different equation must be used to solve for the observer Markov parameters,

$$\bar{\mathbf{Y}} = \mathbf{y} \mathbf{V}_t^T (\mathbf{V}_t \mathbf{V}_t^T)^{-1} \quad (82)$$

where \mathbf{y}_t and \mathbf{V}_t are obtained by deleting the first p columns in \mathbf{y} and \mathbf{V} , respectively.

The system Markov parameters and the observer gain Markov parameters are recovered from the observer Markov parameters using Eq. (28) and Eq. (34), respectively. As mentioned in Section 6.2, application of ERA will produce a realization of both the system state space matrices and the corresponding observer gain.

In the noise-free case, it is sufficient to solve for a minimum number of observer Markov parameters to recover the system Markov parameters. As discussed previously, for a system of order n , this is the minimum value p_{\min} such that $qp_{\min} \geq n$, where q is the number of outputs. However, if more than the minimum number of observer Markov parameters are to be solved for, $p > p_{\min}$, then the matrix \mathbf{V} (or \mathbf{V}_t) is row rank deficient. This implies that the observer Markov parameters are not unique. A solution can still be found by simply replacing the inverse of the quantity in the parentheses in Eq. (81) by a pseudo-inverse.

With noises, however, the matrix \mathbf{V} (or \mathbf{V}_t) tends to be full rank. In this case, the solution given in Eq. (81) or Eq. (82) is the least-squares solution that minimizes the residual

$$\mathbf{e} = \mathbf{y} - \bar{\mathbf{Y}}\mathbf{V} \quad \text{or} \quad \mathbf{e}_t = \mathbf{y}_t - \bar{\mathbf{Y}}\mathbf{V}_t$$

depending on either Eq. (81) or Eq. (82) is used to solve for $\bar{\mathbf{Y}}$. In the presence of process and measurement noises that are white, zero-mean, and uncorrelated with the data, the least squares result can be interpreted in terms of the Kalman filter. The relationship between the deterministic and stochastic approaches is established in Ref. 10. Similar insights can be obtained by examining the structures of the stochastic models considered in Ref. 11, and in Section 8 of this paper.

7.8 Observer Markov Parameters from General Input-Output Data with Prescribed Observer Poles

The problem of identification of observer Markov parameters with prescribed observer poles is briefly presented here. For an observable system, it is always possible to assign the observer poles in any (symmetric) configuration. For simplicity, consider the case where all prescribed poles are real, distinct, with magnitudes less than one. They are denoted by λ_i , $i = 1, 2, \dots, n$. It can be shown that Eq. (21) can be written as

$$y(k) = \alpha \sum_{i=0}^{p-1} \lambda_m^{(\tau)} u(k-i-1) + \beta \sum_{\tau=0}^{p-1} \lambda_q^{(\tau)} y(k-i-1) + Du(k) = [\alpha \quad \beta \quad D] \begin{bmatrix} \phi(k-1) \\ \varphi(k-1) \\ u(k) \end{bmatrix} \quad (83)$$

The quantities $\phi(k-1)$ and $\varphi(k-1)$ are the weighted input and output data vectors

$$\phi(k-1) = \mathfrak{Z}_m \underline{u}(k-p) \quad , \quad \varphi(k-1) = \mathfrak{Z}_q \underline{y}(k-p) \quad (84)$$

where \mathfrak{Z}_m , \mathfrak{Z}_q are the weighting matrices formed by the prescribed eigenvalues,

$$\mathfrak{S}_\ell = \begin{bmatrix} \underline{\lambda}_{1,\ell}^{(p-1)} & \underline{\lambda}_{1,\ell}^{(p-2)} & \cdots & \underline{\lambda}_{1,\ell}^{(1)} & I_{\ell \times \ell} \\ \underline{\lambda}_{2,\ell}^{(p-1)} & \underline{\lambda}_{2,\ell}^{(p-2)} & \cdots & \underline{\lambda}_{2,\ell}^{(1)} & I_{\ell \times \ell} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \underline{\lambda}_{n,\ell}^{(p-1)} & \underline{\lambda}_{n,\ell}^{(p-2)} & \cdots & \underline{\lambda}_{n,\ell}^{(1)} & I_{\ell \times \ell} \end{bmatrix}, \quad \ell = m, q$$

$$\underline{\lambda}_{i,\ell}^{(k)} = \text{diag}(\lambda_i^k, \lambda_i^k, \dots, \lambda_i^k)_{\ell \times \ell}, \quad i = 1, 2, \dots, n, \quad k = 1, 2, \dots, p-1$$

and $\underline{u}(i-p)$ and $\underline{y}(i-p)$ denote the $mp \times 1$ and $qp \times 1$ input and output history vectors defined as

$$\underline{u}(k-p) = \begin{bmatrix} u(k-p) \\ \vdots \\ u(k-2) \\ u(k-1) \end{bmatrix}, \quad \underline{y}(k-p) = \begin{bmatrix} y(k-p) \\ \vdots \\ y(k-2) \\ y(k-1) \end{bmatrix}$$

For convenience, Eq. (83) can be simply written as

$$y(k) = \Upsilon \Gamma(k-1) \quad (85)$$

where

$$\Upsilon = [\alpha \quad \beta \quad D], \quad \Gamma(k-1) = \begin{bmatrix} \phi(k-1) \\ \varphi(k-1) \\ u(k) \end{bmatrix} \quad (86)$$

For a set of $N+1$ measurements of $y(i)$ and $u(i)$, $i = 0, 1, \dots, N$, assuming zero initial conditions, the input-output relation can be written as

$$\mathbf{y} = \Upsilon \mathbf{\Gamma} \quad (87)$$

where \mathbf{y} is given as in Eq. (75) and $\mathbf{\Gamma}$ is

$$\mathbf{y} = [y(0) \quad y(1) \quad \cdots \quad y(p) \quad y(p+1) \quad \cdots \quad y(N)] \quad (88)$$

$$\mathbf{\Gamma} = [\Gamma(-1) \quad \Gamma(0) \quad \cdots \quad \Gamma(p-1) \quad \Gamma(p) \quad \cdots \quad \Gamma(N-1)]$$

with $u(k)$, $y(k)$ set to zero for $k < 0$. The least squares solution for the observer parameters Υ are computed from

$$\Upsilon = \mathbf{y} \mathbf{\Gamma}^T (\mathbf{\Gamma} \mathbf{\Gamma}^T)^{-1} \quad (89)$$

As before, if the initial conditions are not zero then the observer parameters can still be solved from Eq. (89) with \mathbf{y} and $\mathbf{\Gamma}$ replaced by \mathbf{y}_r and $\mathbf{\Gamma}_r$, which are obtained by deleting the first p columns

of \mathbf{y} and $\mathbf{\Gamma}$, respectively. The observer Markov parameters are computed from the observer parameters as

$$\begin{aligned}\bar{\mathbf{Y}}(k) &= C(A + MC)^{k-1} [B + MD \quad -M] \\ &= [\alpha \lambda_m^{(k-1)} \quad \beta \lambda_q^{(k-1)}]\end{aligned}\quad (90)$$

where

$$\lambda_\ell^{(k-1)} = [\underline{\lambda}_{1,\ell}^{(k-1)} \quad \underline{\lambda}_{2,\ell}^{(k-1)} \quad \dots \quad \underline{\lambda}_{n,\ell}^{(k-1)}], \quad \ell = m, q$$

As before, the system Markov parameters and the observer gain Markov parameters are recovered from the observer Markov parameters using Eq. (28) and Eq. (34), respectively. The general case where the prescribed poles are complex or deadbeat can be treated similarly. The readers are referred to Refs. 6 and 8 for further details.

8. Stochastic Linear Difference Equations

This section examines some particular stochastic linear difference equations for the identification of state space models in the presence of noises. In particular, two models are considered that can be interpreted in terms of a Kalman filter.⁹⁻¹¹

8.1 Kalman Filter Model

Consider the case where the state-space equations given in Eq. (8) are extended to include process and measurement noises

$$\begin{aligned}x(k+1) &= Ax(k) + Bu(k) + w_1(k) \\ y(k) &= Cx(k) + w_2(k)\end{aligned}\quad (91)$$

The process noise $w_1(k)$ and measurement noise $w_2(k)$ are two statistically independent, zero-mean, stationary white noise processes with covariances Q and R respectively. The same system can also be expressed in the form of a Kalman filter

$$\begin{aligned}\hat{x}(k+1) &= A\hat{x}(k) + Bu(k) + K\varepsilon(k) \\ y(k) &= C\hat{x}(k) + Du(k) + \varepsilon(k)\end{aligned}\quad (92)$$

where $\varepsilon(k)$ is a white sequence of residual with covariance $\Sigma = CPC^T + R$, and P is the unique positive definite symmetric solution of the algebraic Riccati equation

$$P = APA^T - APC^T(CPC^T + R)^{-1}CPA^T + Q$$

The Kalman filter gain K is given by $K = APC^T(CPC^T + R)^{-1} = APC^T\Sigma^{-1}$. The system in Eq. (91) can also be expressed as

$$\begin{aligned}\hat{x}(k+1) &= (A - KC)\hat{x}(k) + (B - KD)u(k) + Ky(k) \\ y(k) &= C\hat{x}(k) + Du(k) + \varepsilon(k)\end{aligned}\quad (93)$$

Define

$$\tilde{A} = A - KC, \quad \tilde{B} = [B - KD, K], \quad v(k) = \begin{bmatrix} u(k) \\ y(k) \end{bmatrix} \quad (94)$$

Equation (93) becomes

$$\begin{aligned}\hat{x}(k+1) &= \tilde{A}\hat{x}(k) + \tilde{B}v(k) \\ y(k) &= C\hat{x}(k) + Du(k) + \varepsilon(k)\end{aligned}\quad (95)$$

Provided that p_1 is sufficiently large such that

$$\tilde{A}^k = (A - KC)^k \approx 0 \quad k \geq p_1 \quad (96)$$

the input-output description with zero initial conditions can be approximated by

$$y(k) - \sum_{i=1}^{p_1} C(A - KC)^{i-1} Ky(k-i) = \sum_{i=1}^{p_1} C(A - KC)^{i-1} (B - KD)u(k-i) + Du(k) + \varepsilon(k) \quad (97)$$

This equation has the following linear difference structure

$$y(k) + \sum_{i=1}^{p_1} \tilde{A}_i y(k-i) = \sum_{i=0}^{p_1} \tilde{B}_i u(k-i) + \varepsilon(k) \quad (98)$$

where

$$\tilde{A}_i = -C(A - KC)^{i-1} K, \quad \tilde{B}_0 = D \quad (99)$$

$$\tilde{B}_i = C(A - KC)^{i-1} K, \quad i = 1, 2, \dots, p_1$$

Observe that this is an extension of the deterministic linear difference equation presented in Eq. (1). In Eq. (98), the term $\varepsilon(k)$ tends to the residual of the Kalman filter provided that p_1 is sufficiently large such that the condition given in Eq. (96) holds. The stochastic model structure in Eq. (98), which is only valid if p_1 is sufficiently large, has an important interpretation in the deterministic case which does not require p_1 to be as large. This can be seen by re-examining the deterministic formulation in light of the stochastic consideration here. Comparing Eq. (97) to Eq. (21) reveals that $p_1 = p$ and $K = -M$, provided that $\varepsilon(k) \equiv 0$. Thus the matrix K plays the role of $-M$ in the deterministic case. The role of the matrix M in the deterministic case has already been explained. One thus sees that the stochastic equation considered in Eq. (97) takes a precise meaning when reduced to the deterministic case. Consequently, the deterministic solution for the observer Markov parameters given in Eq. (81) or Eq. (82) has a precise interpretation with respect to the Kalman filter when applied to noisy data. This analysis only serves to interpret the meaning of the

residual term in Eq. (98). In system identification, in order to assure that the identified residual indeed converges to the Kalman filter residual, the data set used must be sufficiently long.

The Markov parameters associated with a Kalman filter are called *Kalman filter Markov parameters*. Identification of these parameters recovers not only the system but also an optimal Kalman filter. An important advantage of the structure in Eq. (97) is that the residual $\varepsilon(k)$ appears as an additive term. However, the model structure is valid in describing the input-output relation of the original system given in Eq. (91) in term of the Kalman filter only when p_1 is sufficiently large such that the condition given in Eq. (96) holds. Typically p_1 is required to be several times larger than the order n of the system, $p_1 \gg n$.

8.2 ARMAX Model via a Kalman Filter

In this section, a more general stochastic linear difference model is considered. This model has the form

$$y(k) + \sum_{i=1}^{p_2} A_i y(k-i) = \sum_{i=0}^{p_2} B_i u(k-i) + \sum_{i=0}^{p_2} C_i \varepsilon(k-i) \quad (100)$$

This stochastic linear difference model, also known as an Auto-Regressive Moving Average with eXogeneous input model (ARMAX), is widely used in adaptive estimation, filtering, and control literature. An interpretation of the coefficients of this model in terms of the Kalman filter is given below.

Add and subtract the term $My(k)$ to the right hand side of the state equation in Eq. (92),

$$\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + Bu(k) + My(k) - My(k) + K\varepsilon(k) \\ &= A\hat{x}(k) + Bu(k) + M[C\hat{x}(k) + Du(k) + \varepsilon(k)] - My(k) + K\varepsilon(k) \\ &= (A + MC)\hat{x}(k) + (B + MD)u(k) - My(k) + (M + K)\varepsilon(k) \end{aligned} \quad (101)$$

Define

$$\bar{A} = A + MC, \quad \bar{B} = [B + MD, -M], \quad v(k) = \begin{bmatrix} u(k) \\ y(k) \end{bmatrix}, \quad \bar{M} = M + K \quad (102)$$

the original system in Eq. (91) can be described in the form

$$\begin{aligned} \hat{x}(k+1) &= \bar{A}\hat{x}(k) + \bar{B}v(k) + \bar{M}\varepsilon(k) \\ y(k) &= C\hat{x}(k) + Du(k) + \varepsilon(k) \end{aligned} \quad (103)$$

The input-output description of the above system with zero initial condition is

$$y(k) + \sum_{i=1}^{p_2} C(A + MC)^{i-1} My(k-i) = \sum_{i=1}^{p_2} C(A + MC)^{i-1} (B + MD)u(k-i) + Du(k) + \sum_{i=1}^{p_2} C(A + MC)^{i-1} (M + K)\varepsilon(k-i) + \varepsilon(k) \quad (104)$$

where M and p_2 are such that

$$\bar{A}^k = (A + MC)^k \approx 0, \quad k \geq p_2 \quad (105)$$

For an observable system of order n , it is possible that $p_2 \leq n \ll p_1$. The coefficients of this ARMAX model in Eq. (100) given in terms of A, B, C, M , and K are

$$A_i = C(A + MC)^{i-1} M, \quad i = 1, 2, \dots, p_2$$

$$B_0 = D, \quad B_i = C(A + MC)^{i-1} M \quad (106)$$

$$C_0 = I, \quad C_i = C(A + MC)^{i-1} (M + K)$$

This model has the advantage that p_2 need not be as large as p_1 in the previous model. However, this advantage is off-set by the introduction of the noise dynamics described by the C_i coefficients. This makes the identification problem become non-linear in the parameters since both the C_i coefficients and the residual $\varepsilon(k)$ are not known from the beginning. Note that in this description, the matrix M is used to limit the number of coefficients A_i, B_i, C_i in Eq. (100) due to the term $A + MC$, including the case where $A + MC$ is deadbeat. The matrix M does not play the role of an observer gain for the system given in Eq. (91), which is now given by K . The algebraic relations between the coefficients A_i, B_i of the ARMAX model given in Eq. (106) and the system Markov parameters $Y(k) = CA^{k-1}B$, the observer gain Markov parameters $Y_M(k) = CA^{k-1}M$ are the same as given in Eq. (28) and Eq. (34). Furthermore, the Kalman filter gain Markov parameters $Y_K(k) = CA^{k-1}K$ can also be computed from

$$Y_K(k) = \bar{Y}^{(2)}(k) + \sum_{i=1}^{k-1} \bar{Y}^{(2)}(i) Y_K(k-i) + C_k \quad (107)$$

where $\bar{Y}^{(2)}(k) \equiv 0$ for $k > p$. Application of a realization procedure to the combined sequence of Markov parameters

$$P(k) = \begin{bmatrix} CA^{k-1}B & CA^{k-1}M & CA^{k-1}K \end{bmatrix}$$

$$= CA^{k-1} \begin{bmatrix} B & M & K \end{bmatrix}, \quad k = 1, 2, \dots \quad (108)$$

will realize simultaneously the matrices A, B, C, M , and K .

As a final note, compare the ARMAX model in Eq. (100) with the model considered in Eq. (98). If the model in Eq. (98) is used, but p_1 is not sufficiently large, then the residual in the

model will be colored. In fact, this colored residual can be modelled by the the C_i coefficients in the ARMAX model given in Eq. (106).

9. Concluding Remarks

This paper presents an overview of Markov parameters with particular emphasis on their roles in system identification theory. In discrete-time analysis, the Markov parameter sequences have an important physical interpretation as they are the unit pulse response functions of a linear system. The close relationship between the Markov parameters of a system and its various input-output representations is a basis in system identification theory. In fact, a realization of a linear time-invariant system can be defined as a factorization of a sequence of Markov parameters into a specifically prescribed form. When driven by certain specialized inputs, the system sampled response can be conveniently interpreted as Markov parameters. Recent results in system identification from general inputs generated renewed interests in Markov parameters. In particular, the Markov parameters can be extended to include observer Markov parameters and Kalman filter Markov parameters. These concepts have been found to be quite useful in developing identification techniques for application in complex flexible structures.

10. References

- 1 Kailath, T., *Linear Systems*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1980.
- 2 Ljung, L. and Söderström, T., *Theory and Practice of Recursive Identification*, The MIT Press, Cambridge, Massachusetts, 1983.
- 3 Chen, C.-T., *Linear System Theory and Design*, CBS College Publishing, New York, New York, 1984.
- 4 Goodwin, G.C. and Sin, K.S., *Adaptive Filtering, Prediction, and Control*, Prentice Hall, Englewood Cliffs, New Jersey, 1984.
- 5 Phan, M., Juang, J.-N., and Longman, R.W., "Identification of Linear Multivariable Systems from a Single Set of Data by Identification of Observers with Assigned Real Eigenvalues," Paper No. 91-0949, *Proceedings of the 32nd Structures, Structural Dynamics, and Materials (SDM) Conference*, Baltimore, Maryland, April 1991.
- 6 Phan, M., Horta, L.G., Juang, J.-N., and Longman, R.W., "Linear System Identification Via An Asymptotically Stable Observer," Paper No. AIAA-91-2734, *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, New Orleans, Louisiana, August 1991.
- 7 Horta, L.G., Phan, M., Juang, J.-N., Longman, R.W., and Sulla, J.L., "Frequency Weighted System Identification and Linear Quadratic Controller Design," Paper No. AIAA-91-2733, *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, New Orleans, Louisiana, August 1991.
- 8 Phan, M., Horta, L.G., Juang, J.-N., and Longman, R.W., "Identification of Linear Systems Via An Asymptotically Stable Observer," *NASA Technical Paper 3164*, 1992.

- 9 Chen, C.-W., Huang, J.-K., Phan, M., and Juang, J.-N., "Integrated System Identification and Modal State Estimation for Control of Large Flexible Space Structures," Paper No. 90-3740-CP, *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Portland, Oregon, August 1990.
- 10 Juang, J.-N., Phan, M., Horta, L.G., and Longman, R.W., "Identification of Observer/Kalman Filter Markov Parameters: Theory and Experiments," Paper No. AIAA-91-2735, *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, New Orleans, Louisiana, August 1991.
- 11 Phan, M., Horta, L.G., Juang, J.-N., and Longman, R.W., "Improvement of Observer/Kalman Filter Identification by Residual Whitening," *Proceedings of the 1991 VPI&SU Symposium on Dynamics and Control of Large Structures*, Blacksburg, Virginia, May 1991.
- 12 Ho, B.L. and Kalman, R.E., "Effective Construction of Linear State Variable Models from Input/Output Data," *Proceedings of the 3rd Annual Allerton Conference on Circuits and Systems Theory*, Monticello, Illinois, 1965.
- 13 Juang, J.-N. and Pappa, R.S., "An Eigensystem Realization Algorithm for Modal Parameter Identification and Model Reduction," *Journal of Guidance, Control, and Dynamics*, Vol. 8, September-October 1985.
- 14 Juang, J.-N., Cooper, J.E., and Wright, J.R., "An Eigensystem Realization Algorithm Using Data Correlations (ERA/DC) for Modal Parameter Identification," *Control Theory and Advanced Technology*, Vol. 4, No. 1, 1988.
- 15 Bergmann, M., *Enhanced System Identification for Large Flexible Spacecraft*, Ph.D. Dissertation, Columbia University, New York, New York, 1990.
- 16 Akaike, H., "Stochastic Theory of Minimal Realization," *IEEE Transactions on Automatic Control*, Vol. AC-19, No. 6, December 1974.
- 17 Juang, J.-N., Horta, L.G., Phan, M., and Longman, R.W., "Learning System Identification for Flexible Structures," *Proceedings of the 2nd USAF/NASA Workshop on System Identification and Health Monitoring of Precision Space Structures*, California Institute of Technology, Pasadena, California, March 1990.
- 18 Phan, M., Juang, J.-N., and Longman, R.W., "Recent Developments in Learning Control and System Identification for Robots and Structures," *Dynamics of Flexible Structures in Space*, C.L. Kirk and J.L. Junkins, Editors, Computational Mechanics Publications, Southampton, Boston and Springer-Verlag, New York, co-publishers, May 1990.
- 19 Van Loan, C., "Computing Integrals Involving the Matrix Exponential," *IEEE Transactions on Automatic Control*, Vol. AC-23, No. 3, June 1978.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE

October 1991

3. REPORT TYPE AND DATES COVERED

Technical Memorandum

4. TITLE AND SUBTITLE

On Markov Parameters in System Identification

5. FUNDING NUMBERS

WU 590-14-21-01

6. AUTHOR(S)

Minh Phan, Jer-Nan Juang, and Richard W. Longman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

NASA Langley Research Center
Hampton, Virginia 23665-5225

8. PERFORMING ORGANIZATION
REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING/MONITORING
AGENCY REPORT NUMBER

NASA TM-104156

11. SUPPLEMENTARY NOTES

Minh Phan-Lockheed Engineering and Sciences Co., Hampton, VA; Jer-Nan Juang-NASA Langley Research Center, Hampton, VA; Richard Longman-Columbia University, New York, New York

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified - Unlimited

Subject Category 39

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This paper presents a detailed discussion of Markov parameters in system identification. Different forms of input-output representation of linear discrete-time systems are reviewed and discussed. Interpretation of sampled response data as Markov parameters is presented. Relations between the state-space model and particular linear difference models via the Markov parameters are formulated. A generalization of Markov parameters to observer and Kalman filter Markov parameters for system representation and identification is explained. These extended Markov parameters play an important role in providing not only a state-space realization, but also an observer/Kalman filter for the system of interest.

14. SUBJECT TERMS

Markov Parameters, Observer Markov Parameters, Observer/Kalman Filter Identification (OKID), State Space Realization

15. NUMBER OF PAGES

31

16. PRICE CODE

A03

17. SECURITY CLASSIFICATION
OF REPORT

Unclassified

18. SECURITY CLASSIFICATION
OF THIS PAGE

Unclassified

19. SECURITY CLASSIFICATION
OF ABSTRACT

Unclassified

20. LIMITATION OF ABSTRACT